Investigating the FOOT-STRUT distinction in Northern Englishes using crowdsourced data

Strycharczuk, Patjcia ; Brown, Georgina ; Leemann, Adrian ; Britain, David

Abstract: The FOOT and STRUT lexical sets did not undergo a historical split in the North of England, and these vowels are said to remain a single phoneme for present day Northern English speakers. However, several sources report variation in this respect. We analyse this variation in production, using acoustic analysis of crowdsourced data from 141 speakers of seven Northern English urban dialects. 36 speakers in our sample show a categorical distinction between FOOT and STRUT. Highly mobile speakers are more likely to have this distinction, compared to speakers with low mobility. A categorical split is also more likely in speakers from Newcastle, compared to several other cities. While we find no evidence that FOOT and STRUT vowels are splitting in the North, we discuss how the observed variation may contribute to the presence of marginal contrasts.

Originally published at:
Strycharczuk, Patjcia; Brown, Georgina; Leemann, Adrian; Britain, David (2019). Investigating the FOOT-STRUT distinction in Northern Englishes using crowdsourced data. In: 19th International Congress of Phonetic Sciences, Melbourne, Australia, 5 August 2019 - 9 August 2019. Australasian Speech Science and Technology Association Inc, 1337-1341.

INVESTIGATING THE FOOT-STRUT DISTINCTION IN NORTHERN ENGLISHES USING CROWDSOURCED DATA

Patrycja Strycharczuk[1], Georgina Brown[2], Adrian Leemann[2,3] and David Britain[3]

[1]University of Manchester, [2]Lancaster University, [3]University of Bern
patrycja.strycharczuk@manchester.ac.uk, g.brown5@lancaster.ac.uk, adrian.leemann@csls.unibe.ch, britain@ens.unibe.ch

ABSTRACT

The FOOT and STRUT lexical sets did not undergo a historical split in the North of England, and these vowels are said to remain a single phoneme for present day Northern English speakers. However, several sources report variation in this respect. We analyse this variation in production, using acoustic analysis of crowdsourced data from 141 speakers of seven Northern English urban dialects. 36 speakers in our sample show a categorical distinction between FOOT and STRUT. Highly mobile speakers are more likely to have this distinction, compared to speakers with low mobility. A categorical split is also more likely in speakers from Newcastle, compared to several other cities. While we find no evidence that FOOT and STRUT vowels are splitting in the North, we discuss how the observed variation may contribute to the presence of marginal contrasts.

Keywords: Northern English; FOOT-STRUT contrast; vowel split; crowdsourced data

1. INTRODUCTION

The absence of a phonemic contrast between FOOT and STRUT vowels is traditionally a major defining feature for Northern varieties of British English [2, 19]. It is also a salient marker of Northern English, sometimes taken as an isogloss demarcating the North of England from the South [18, 19].

However, several sources note that the FOOT and STRUT vowels may be distinct for some Northern English speakers [1, 3, 19], as well as for East-Midlands speakers who traditionally do not have a split [5]. In a study of 123 speakers from Greater Manchester, Baranowski & Turton find eight who have a phonemic distinction, i.e. FOOT and STRUT are separate phonemic categories in production and perception, similar to Southern English speakers. Furthermore, some speakers in the same study who do not have a phonemic contrast, may show small but systematic phonetic differences between FOOT and STRUT, mainly realised as F1 lowering for the STRUT category.

In this paper, we investigate whether systematic production distinctions between FOOT and STRUT vowels are also present across a selection of urban Northern English accents spoken in Hull, Leeds, Liverpool, Manchester, Newcastle upon Tyne, Sheffield and York. Our interest is in the overall distribution of the phonetic distance between these two categories, and whether such distribution is constrained by any geographic or social factors.

Our data come from the English Dialects App Corpus (EDAC) [13, 14]. They are crowdsourced recordings of the passage 'The Boy who Cried Wolf', collected via mobile phones. The corpus currently contains recordings from 3,500 speakers in the British Isles (including Republic of Ireland), and as such, it is an extremely rich resource for studying dialectal variation. The text contains all vowels of English, and hence it is useful for mapping out vowel spaces [8]. From the technical point of view, we expect the recordings to be generally suitable for an instrumental analysis of formant values [7].

At the same time, the corpus was not designed to study any particular phonetic variable. This is an important factor for us to consider, because we want to determine whether two vowels are merged or not, and established ways of doing that are not applicable to this type of data. The gold standard in determining the presence of a phonemic distinction is the minimal pairs test in production and perception, but there are no FOOT-STRUT minimal pairs in the corpus, nor do we have perception data. The second problem is that the text is quite short, and therefore relatively few words from each lexical set are present (see Table 1). Consequently, any attempts to apply by-speaker statistical measures, such as Pillai-scores [11], raise concerns of statistical power. Finally, the effects of lexical set and phonological environment are confounded in the text. For instance, there are only four distinct FOOT words, and two of those involve a following /l/, which is known to exert a strong F2 lowering effect on the preceding /ʊ/ [17]. Out of the eight STRUT words, four have a

nasal following the vowel, which is likely to raise the F1 [12]. Prosodic factors, such as size of the word, are not controlled for either.

Table 1: Items from the FOOT and STRUT lexical sets in 'The Boy who Cried Wolf' passage

FOOT	STRUT
full; foot; good; wolf (4x)	*come; company; cousins; duck; fun; much; some; rushed*

The problem with the phonological environment confounds is that they systematically affect the distances we can measure between the two vowels: We cannot estimate the distance between the two vowels independently of the phonological environment. Nycz and Hall-Lew discuss this issue in the context of near-mergers [16]. Although our case does not technically involve a merger, it bears some phonetic similarities to a near-merger situation, since previous reports indicate that FOOT and STRUT classes largely overlap in the North, but small systematic differences between them still persist [1].

Due to the nature of EDAC data, and conscious of the limitations discussed above, we investigate the distance between FOOT and STRUT vowels from the point of view of inter-speaker variation. While we may not be able to determine for each and every speaker whether the two vowels form a single class or not, we are able to systematically compare the FOOT and STRUT realisation between speakers, because all speakers read the same text. Based on this, we investigate the range of distances between FOOT and STRUT categories in our sample of Northern speakers, and we identify subpopulations of speakers, based on Gaussian mixture modelling.

2. MATERIALS & METHOD

2.1. Materials

As discussed in Section 1 above, EDAC is based on recordings of 'The Boy who Cried Wolf' passage. In this passage, we focused on the FOOT and STRUT words listed in Table 1. We also included measurements of three items including vowels closer to the corners of the vowel space: *feast* (lexical set of FLEECE), *plan* (lexical set of TRAP) and *thought*.

For a detailed description of how the original data were acquired, see [14].

2.2. Speakers

For the present analysis, we selected 141 speakers from EDAC. They represent seven cities in the North of England: Hull (N =15), Leeds (N =27), Liverpool (N =19), Manchester (N =25), Newcastle upon Tyne (N =22), Sheffield (N =19) and York (N = 14). One of the reasons for selecting these cities was geographic diversity: the cities are spread over a relatively large area, and they represent a variety of Northern urban dialects. Our rationale for focusing on urban varieties is that we expect urban dialects to show more variation, due to migration patterns and increased contact with non-local dialects, which, in the British English context, may involve increased exposure to varieties with the FOOT and STRUT split.

The main criterion in the selection of particular cities was the number of speakers in the corpus representing each city. For each city, we selected all the speakers for whom we had the complete passage recording, also checking that the recording quality was clear enough for instrumental analysis. We made sure that there were no interruptions, no additional talkers present, and no excessive background noise. The speakers were asked to determine themselves which dialect they represent, by putting a pin on a locality that best corresponds to their dialect.

The age of the speakers ranged from 9 to 73 years, with a median of 28. 50% of the participants were between 21 and 36 years of age. The sample included 82 female and 61 male speakers.

Additional information we have about the speakers concerns mobility, operationalised as the number of times the person has moved over the last 10 years. There were four response categories, corresponding to 0, 1, 2–3, and 4 moves or more. The responses in our sample were evenly distributed between the four groups. with 35, 32, 39 and 32 speakers in each respective category.

2.3. Data processing

The audio files were forced aligned using an HTK-based forced aligner that was developed in-house. For some of the tokens, the onset and offset of the vowel were manually corrected by two undergraduate research assistants. This was the case for all the tokens of *duck*, *foot*, *feast*, *plan*, *thought*. For the remaining tokens, we proceeded with the output of the forced alignment.

We measured the first two formants of each vowel token at the acoustic midpoint, using a Praat script. We used the Burg algorithm in Praat [4], with the standard settings (maximum 5 formants, 25ms window, 50 Hz pre-emphasis). For male speakers, we

set the maximum formant at 5kHz, whereas for females speakers, it was 5.5kHz.

Since the automated labelling and measurement procedure inevitably produces some errors, we removed outliers, defined as values removed more than 2.5 SDs from the mean value of F1 and F2 within each lexical set for each speaker. We then z-score normalised the F1 and F2 values within each speaker [15] to reduce inter-speaker variation related to anatomical differences. We included the corner vowels in the scaling.

2.4. Statistical analysis

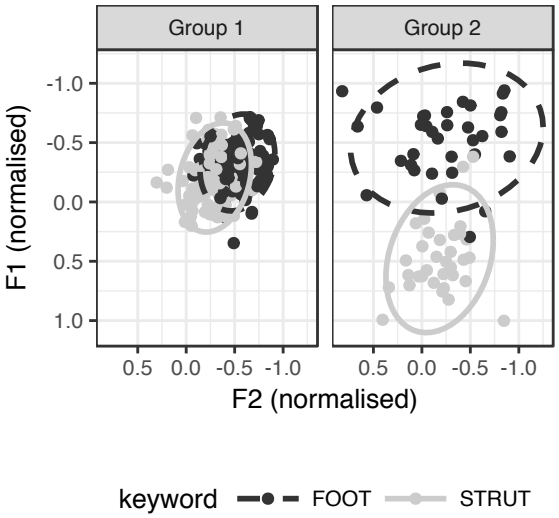
For each speaker, we calculated the median value of F1 and F2 for the FOOT and STRUT sets. In this way, we obtained four distinct values per speaker, which we then used as the input to a Gaussian mixture modelling analysis implemented in `R`, using the `mclust` package [10, 9]. This procedure estimates the probability of the existence of subpopulations within the overall population. We expect that the default realisation of FOOT and STRUT in North of England should involve no contrast between those categories, although we also expect that differences between these two categories may be detectable in our data, due to phonological environment confounds. If some speakers have a categorical FOOT - STRUT distinction in production, we expect them to pattern differently from the remaining speakers, i.e. they will form a separate group.

3. RESULTS

Based on mixture modelling, the best-fitting model of the data had 2 mixture components, which suggests there are two latent groups amongst our speakers. Figure 1 shows the two clusters of speakers that were identified. The two facets correspond to the two groups identified by the model, and each subplot shows the relative realisation of FOOT and STRUT within each group. For group 2, FOOT and STRUT categories are completely separate. This separation is mainly along the F1 dimension, with considerable F1 raising for STRUT, which is consistent with the way the FOOT - STRUT split is realised in the South of England. In addition, group 2 shows more variance along the F2 dimension, with some relatively front median FOOT realisations. Group 2 (categorically separated) comprises 34 speakers (24% of our sample).

For group 1 (107 speakers), the two vowel categories overlap to an extent, although they are not identical. Based on Wilcoxon tests, there was a significant difference in normalised F1 and F2, depending on the lexical set within this speaker group, with $p-$values lower than 0.001 for both normalised F1 and F2.

Figure 1: Results of clustering showing two groups of speakers distinct in their realisation of FOOT and STRUT



In order to gain a better understanding of who the speakers in each group are, we applied a generalised linear model, with the classification (group 1 or group 2) as the dependent variable, and with age, sex, city and mobility as predictors. Predictors were interpreted as significant if p was lower than .05. The results show an effect of mobility, in which highly mobile speakers (those who have moved more than four times in the past years) are more likely to have a categorical split, compared to speakers who have not moved at all in the same time span (β =1.46, SE = 0.70, t =2.084, p <.05). In terms of geographical distribution, Newcastle speakers were significantly more likely to have a categorical split, compared to speakers from Leeds, Manchester, Sheffield and York ($p < .05$ for all pairs). No significant differences were found between any other pairs of cities. Table 2 shows the numbers of speakers from each city classified as group 1 and group 2.

4. DISCUSSION

In our sample of 141 speakers who self-identified as speakers of an urban Northern English accent, almost a quarter (34 speakers) had a considerable distance between FOOT and STRUT lexical sets, such that these two sets form separate categories in pro-

Table 2: The number of speakers in each city classified as group 1 and 2

	Group 1	Group 2 (categorical split)
Hull	12	3
Leeds	22	5
Liverpool	15	4
Manchester	20	5
Newcastle upon Tyne	11	11
Sheffield	16	3
York	11	3

duction at the level of this sub-population. Although in the absence of perception data, we cannot conclude that these speakers have a phonemic FOOT and STRUT distinction, their production pattern is similar to how FOOT and STRUT are realised in the South of England, where they are separate phonemes.

Although we expected to identify speakers like this in our Northern sample, it is quite striking that there are as many. Recall that in a large recent Manchester corpus, only about 6% showed a phonemic FOOT and STRUT contrast [1], whereas we find larger proportions both across the North and in Manchester alone. We speculate that the discrepancy between the two studies can be explained by potential differences in speaker sampling with respect to mobility patterns. Baranowski & Turton surveyed native speakers of Manchester English resident in Manchester. In contrast, EDAC imposes no such restrictions, and thus the respondents may include speakers who have moved away from their native area, but who still identify as speakers of a particular accent, or speakers who have grown up in a particular locality, but whose parents come from elsewhere. Therefore, the frequency of a categorical FOOT - STRUT distinction we find may not be representative of its frequency in more 'pure' Northern English accents.

Imposing restrictions on speaker mobility is a feature of traditional dialectology, inherent to its reliance on NORMs (Non-mobile Old Rural Males). EDAC represents a radical departure from this methodological approach, and hence it is not a suitable tool for identifying traditional accents. On the other hand, the type of variation recorded in EDAC is probably more representative of actual variation found in modern day urban localities, in which high levels of migration lead to increased mixing between different varieties. Taking the example of Manchester, the city's population grew by 19% between 2001 and 2011, reaching 503,127 at the time of 2011 census [6]. Over this period, more than 30,000 people moved into Manchester each year from different locations in the UK, while a slightly higher number left Manchester for elsewhere in the UK.

These trends suggest that highly-mobile individuals may be increasingly the norm, and therefore mixed dialects are becoming an important feature of the overall linguistic landscape. The importance of mobility as a contributing factor to language variation is underscored by our finding that low mobility, operationalised as having lived in the same household for ten years or more, is predictive of having considerable overlap between FOOT and STRUT categories, whereas highly-mobile speakers are more likely to have a categorical split.

One way to interpret the mobility findings is in terms of standardisation and dialect levelling [20]. Since the FOOT-STRUT split is a feature of Standard (Southern) English, the relatively more mobile speakers may be adopting this variant. On the other hand, however, the standardisation account is complicated by the fact that we do not find an effect of other predictors known to affect the adoption of prestige forms, such as gender. Neither do we find evidence that the distinction between FOOT and STRUT is a change in progress. Thus, our current results present a picture, in which the presence of a FOOT-STRUT distinction is a distinct minority pattern, relatively stable across speakers of different age and sex. Furthermore, while Newcastle stands out in having relatively more speakers with a categorical split, such speakers are found in all the cities in our data.

The presence of such variation compels us to ask how it might influence the conservative Northern unsplit realisation of FOOT and STRUT. We note that FOOT and STRUT differ significantly among the F1 and F2 dimensions for those speakers who do *not* have a categorical distinction. This may be due to the unbalanced phonological environment in our test items, as discussed in Section 1, but we could also hypothesise that, in addition to coarticulation effects, the increased presence of unmerged FOOT and STRUT exemplars to which all Northern speakers are exposed may lead to small, but consistent differences in production for some groups of speakers. We could not verify this hypothesis using present data, but it could be tested in a follow-up study including Northern English speakers, and using more controlled test items. If a small distinction persists, it would support the existence of a marginal split, phonetically not unlike a near-merger.

5. ACKNOWLEDGMENTS

We wish to thank the undergraduate research assistants, Enhui Cheng and Chloe Gornall, for their

work on data processing. We also thank ICPhS reviewers for helpful comments and suggestions.

6. REFERENCES

[1] Baranowski, M., Turton, D. 2018. The FOOT-STRUT vowels in Manchester: Evidence for the diachronic precursor to the split? *University of Pennsylvania Working Papers in Linguistics* 24(2), 2.

[2] Beal, J. 2008. English dialects in the North of England: phonology. In: Schneider, K., Edgar W.and Burridge, Kortmann, B., Mesthrie, R., Upton, C., (eds), *A handbook of varieties of English* volume 1: Phonology. Mouton de Gruyter 113–133.

[3] Beal, J. C. 2010. *Introduction to Regional Englishes*. Edinburgh University Press.

[4] Boersma, P., Weenink, D. 2017. Praat: doing phonetics by computer [Computer programme]. Version 6.0.24.

[5] Braber, N., Flynn, N. 2015. The East Midlands: Nottingham. 369–392.

[6] Bullen, E. 2015. Manchester Migration: A profile of Manchester's migration patterns.

[7] De Decker, P., Nycz, J. 2011. For the record: Which digital media can be used for sociophonetic analysis? *University of Pennsylvania Working Papers in Linguistics* 17(2), 7.

[8] Deterding, D. 1997. The formants of monophthong vowels in Standard Southern British English pronunciation. *Journal of the International Phonetic Association* 27, 47–55.

[9] Fraley, C., Raftery, A. E. 2002. Model-based clustering, discriminant analysis and density estimation. *Journal of the American Statistical Association* 97, 611–631.

[10] Fraley, C., Raftery, A. E., Murphy, T. B., Scrucca, L. 2012. *mclust Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation*.

[11] Hall-Lew, L. 2010. Improved representation of variance in measures of vowel merger. *Proceedings of Meetings on Acoustics 159ASA* volume 9 1–10.

[12] House, A. S., Stevens, K. N. 1956. Analog studies of the nasalization of vowels. *Journal of Speech and Hearing Disorders* 21, 218–232.

[13] Leemann, A., Kolly, M.-J., Britain, D. 2016. English Dialect App.

[14] Leemann, A., Kolly, M.-J., Britain, D. 2018. The English Dialects App: The creation of a crowdsourced dialect corpus. *Ampersand* 5, 1–17.

[15] Lobanov, B. M. 1971. Classification of Russian vowels spoken by different speakers. *The Journal of the Acoustical Society of America* 49, 606–608.

[16] Nycz, J., Hall-Lew, L. 2013. Best practices in measuring vowel merger. *Proceedings of Meetings on Acoustics 166ASA* volume 20.

[17] Strycharczuk, P., Scobbie, J. M. 2017. Fronting of Southern British English high-back vowels in articulation and acoustics. *The Journal of the Acoustical Society of America* 142, 322–331.

[18] Trudgill, P. 2000. *The dialects of England*. Oxford: Blackwell.

[19] Wells, J. 1982. *Accents of English* volume 2. Cambridge: Cambridge University Press.

[20] Williams, A., Kerswill, P. 1999. Dialect levelling: continuity vs. change in Milton Keynes, Reading and Hull. In: *Urban voices: Accent studies in the British Isles*. London: Arnold 141–162.